DOCUMENT RESUME

ED 295 478 FL 017 419

AUTHOR Bardovi-Harlig, Kathleen; Bofman, Theodora
TITLE A Second Look at T-Unit Analysis.
PUB DATE Mar 88
NOTE 17p.; Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (22nd, Chicago, IL, March 8-13, 1988).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Secondary Education; Grammar; *Language Tests; *Measurement Techniques; Phrase Structure; Second Language Instruction; Second Language Learning; *Sentence Structure; *Writing Skills
IDENTIFIERS *T Units

ABSTRACT

An evaluation of the t-unit as the basic unit of analysis of second language development in written form argues that despite the t-unit's advantages, the sentence-based analysis is superior for examining syntactic complexity in at least three ways: (1) it better characterizes learner knowledge; (2) it facilitates comparison across learners and across stages of second language development; and (3) it encourages dialogue between classroom teachers and researchers by providing a common unit of analysis. The t-unit completely eliminates full-clause coordination from any quantitative description of syntactic development, while sentence analysis allows such knowledge to be represented. In addition, the coordination index in sentence analysis allows comparison of the shift from the use of coordination by beginning learners to the use of embedding by more advanced learners, and finally, the use of the sentence, a learner-centered unit, gives both teachers and researchers access to data in the same form. It is suggested that whether used to measure average sentence length, average length of error-free sentences, or clauses per sentence, sentence-based analyses are useful in the analysis of written second language discourse. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED295478

A Second Look at T-Unit Analysis

Kathleen Bardovi-Harlig
Indiana University

Theodora Bofman
Northeastern Illinois University

Paper presented at the 22nd Annual TESOL Convention
8-13 March 1988, Chicago, Illinois

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
K. Bardovi-Harlig

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL 017 449

A Second Look at T-Unit Analysis*

Kathleen Bardovi-Harlig
Indiana University

Theodora Bofman
Northeastern Illinois University

The T-unit, which stands for minimal terminable unit, was introduced by Hunt (1965) to measure development of sentences in the writing of grade school children. The T-unit contains one independent clause and its dependent clauses. Hunt described T-units as "the shortest units into which a piece of discourse can be cut without leaving any sentence fragments as residue" (Hunt 1970, p. 189). A sentence is analyzed as two (or more) T-units when two (or more)independent clauses (with subjects abd finite verbs) are conjoined as in (1a), but as a single T-unit when one or more clauses are embedded in an independent clause as in (1b).

(1a) $[S \text{ and } S]_S$ = 2 T-units

(b) $[S\ [S]]_S$ = 1 T-unit

Examples (2) and (3) illustrate cases (1a) and (b), respectively.

(2) There was a woman next door and she was a singer. = 2 T-units

(3) There was a woman next door who was a singer. = 1 T-unit

Although originally designed to assess development in first language acquisition, the T-unit has been widely adopted for analysis in second language acquisition. The T-unit has been used to examine both oral and written second language production in such diverse endeavors as the description of risk-taking (Beebe 1983) and assessment of global proficiency (Larsen-Freeman 1983) as well as to examine linguistic input and teachers' classroom language (Gaies 1977).

The T-unit has two main advantages. First, in native or non-native oral production, sentence boundaries are difficult to identify. Because T-units divide the speech stream into main clauses and their dependent clauses, the

need to identify sentences is eliminated. Second, in first language for which the T-unit was originally designed, the T-unit eliminated cases of excessive coordination. In his 1970 article, "Recent measures in syntactic development," Hunt sought to account for the increase in length of sentences produced by children at three grade levels and by adults. Because children in the fourth grade (Hunt's youngest group) frequently produce what are known in the classroom as run-on sentences--such as that in example (4)--younger children produce longer sentences than older children or than adults.

(4) I like the movie we saw about Moby Dick, the white whale the captain said if you can kill the white whale, Moby Dick, I will give this gold to the one who can do it and it is worth sixteen dollars they tried and tried but while they were trying they killed a whale and used the oil for the lamps they almost caught the white whale. (English L1, 4th grade)

(5) I like the movie we saw about Moby Dick, the white whale [Ø] the captain said if you can kill the white whale, Moby Dick, I will give this gold to the one who can do it and it is worth sixteen dollars [Ø] they tried and tried but while they were trying they killed a whale and used the oil for the lamps [Ø] they almost caught the white whale. (6 T-units/1 sentence)

By dividing sentences such as (4) into T-units as shown in (5), Hunt could account for the observation that children write increasingly longer units, in this case T-units, as they get older.

Although the T-unit has advantages for certain types of analyses, we have found that it also has disadvantages. During the last two years we have been studying syntactic complexity in compositions written by advanced learners of English with TOEFL scores of approximately 550 (Bardovi-Harlig & Bofman, in press). When we evaluated the essays for syntactic complexity using the T-unit we found that we could not capture certain characteristics of the learner-produced language.[1]

We would like to argue that for the analysis of written discourse of advanced second language learners in particular, and more generally for the analysis of written discourse by second language learners at any level, the

sentence is superior to the T-unit as the basic unit of analysis for investigations of syntactic complexity. To account for cases of excessive coordination (such as run-on sentences) for which the T-unit was designed, we will propose a coordination index.

We will argue that on at least three counts a sentence-based analysis is superior to a T-unit when examining the question of syntactic complexity. First, a sentence-based analysis better characterizes learner knowledge; second, it facilitates comparsion across learners and across stages of second language development; and, third, it encourages dialogue between classroom teachers and researchers by providing a common unit of analysis. We will examine each advantage individually in the sections that follow.

Learner knowledge

First, and most important, the T-unit analysis does not accurately reflect the knowledge of the learner. As the unit directly produced by the learner, the sentence has a certain degree of psychological reality in that it allows researchers to glimpse, through the learners's own production, how the learner views the structure of the English sentence. A T-unit analysis divides sentences which were intended to be units by the language learner.

At both high and low levels of development, a T-unit analysis discounts the learner's knowledge of coordination, a recursive feature of language. Since the basic definition of the T-unit as an independent clause divides all full clause conjuncts into separate T-units, it treats all conjoined and non-conjoined sentences equally, as if they were non-conjoined sentences. As a result, the learner's knowledge of coordination is not acknowledged because coordination is not represented in a T-unit analysis. Thus, the T-unit divides learner-produced text into artificially homogeneous units.

While it's true that a T-unit analysis divides a cumbersome run-on sentence such as (4) into manageable pieces, this same division breaks up

legitimately coordinated sentences which are not only grammatical but which reflect a certain rhetorical sophistication which would be absent if the two clauses were merely juxtaposed as they are in (7). (Example (6) and all following examples identified by subject number originated in essays written by advanced second language learners (TOEFL 543-563) enrolled in university courses (Bardovi-Harlig and Bofman, in press).)

(6) Hundreds of schools were built and tens of institutions are starting to join in providing technical education to the public. (Subject 6) (2 T-units/1 sentence)

(7) Hundreds of schools were built. Tens of institutions are starting to join in providing technical education to the public. (2 T-units/2 sentences)

Not only does the T-unit analysis divide sentences exhibiting coordination which seems to have an additive function as in (6), it also divides sentences in which the conjunction does not have an additive function. In these cases the separated or individual conjuncts do not have the same meaning as the conjoined unit. Take cases of asymetrical conjunction (Lakoff 1971, Schmerling 1976), for instance, in which the conjuncts cannot be reversed. An example of asymetrical conjunction can be found in (8) in which the conjoined senetnces together function as a conditional sentence (Quirk, Greenbaum, Leech & Svartvik 1985, p. 832, 931) similar to the conditional in (9).

(8) Take notice on his characteristics and you'll find that there is different between a Chinese and an Asian from other nations. (Subject 11) (2 T-units/1 sentence)

(9) [If you] take notice of his characteristics then you'll find that there is a difference between a Chinese and an Asian from other nations. (1 T-unit/1 sentence)

By using the sentence as the unit of analysis, examples (8) and (9) receive the same analysis; if T-units were used, then (8) would have two T-units, but (9) would have only one. In other words, a sentence anaylsis reflects the complexity of a conjunctive-conditional such as (8), but a T-unit analysis

does not.

Furthermore, in dividing conjuncts the T-unit analysis treats all conjunctions as semantically null, breaking up not only syntactic units, but semantic units as well. Just as a T-unit analysis divides coordinated sentences joined by and, it also divides conjuncts joined by but and or into separate units. Analyzing sentences such as (10) into two units divides a semantic whole; the contrast between the first and second clause is part of the information conveyed by the sentence.

(10) They do not encourage their children to mix with the opposite sex, but they do let them interact to certain extent. (Subject 19) (2 T-units / 1 sentence)

The T-unit analysis artificially divides sentences which learners see as syntactic units, imposing uniformity of length and complexity on output which is not present in the original production by restricting units to single main clauses.

If the T-unit analysis seems to underestimate a learner's knowledge of coordination on the one hand, on the other it can also give the learner too much credit by breaking up sentences. The sentence in example (11) is overly complex and cumbersome, yet a T-unit analysis misses its awkwardness by separating the conjuncts.

(11) When you see a group of Asians talking, if they are speaking their native language, I think you do not know which native they come from, but if you hear that among their conversation there are some English single words between their native language, I can say that they are Chinese from Hong Kong. (Subject 11) (2 T-units/ 1 sentence)

A sentence analysis captures the full effect of this overloaded sentence whereas a T-unit analysis, by breaking the sentence into two T-units, reduces any measure of complexity by one-half. As an illustration, let us consider two measurements of complexity, the subordinate clause index (number of clauses/syntactic unit) and average length (number of words/syntactic unit) (Hunt 1970). For example, a sentence analysis of this particular sentence

would count 10 clauses/sentence or 53 words/sentence, while a T-unit analysis would count an average of 5 clauses/T-unit (6 clauses/T-unit and 4 clauses/T-unit, respectively) or 26.5 words/T-unit (26 words/T-unit and 27 words/T-unit, respectively). Thus, in this case a T-unit analysis quantitatively reduces the excessive sentence combining on the part of the learner, obscuring the inappropriateness of so many combinations.

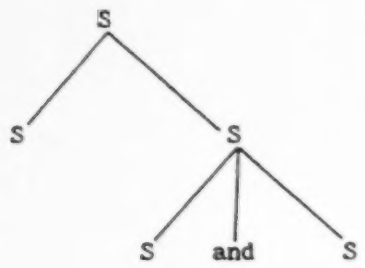
So far we have identified two problems with the T-unit analysis. By reducing sentences into T-units on one hand, knowledge and correct use of coordination is discounted, while on the other, the division of sentences into T-units simplifies sentences which otherwise seem to violate limits on complexity or information giving them the appearance (at least quantitatively) of being of average length and complexity.

A third difficulty lies in the differential treatment of full clause coordination depending on the environment. When the conjuncts are main clauses as in (12), a T-unit analysis divides the conjuncts as we have seen in example (1a). But when the conjuncts are subordinate clauses with the configuration shown in (13), exemplified by (14), no division is made.[2] As a result, a T-unit account treats coordination in two different ways although the coordination of any two full clauses presumably involves the same linguistic knowledge and in formal linguistic accounts receives the same description.

(12) Main clause coordination

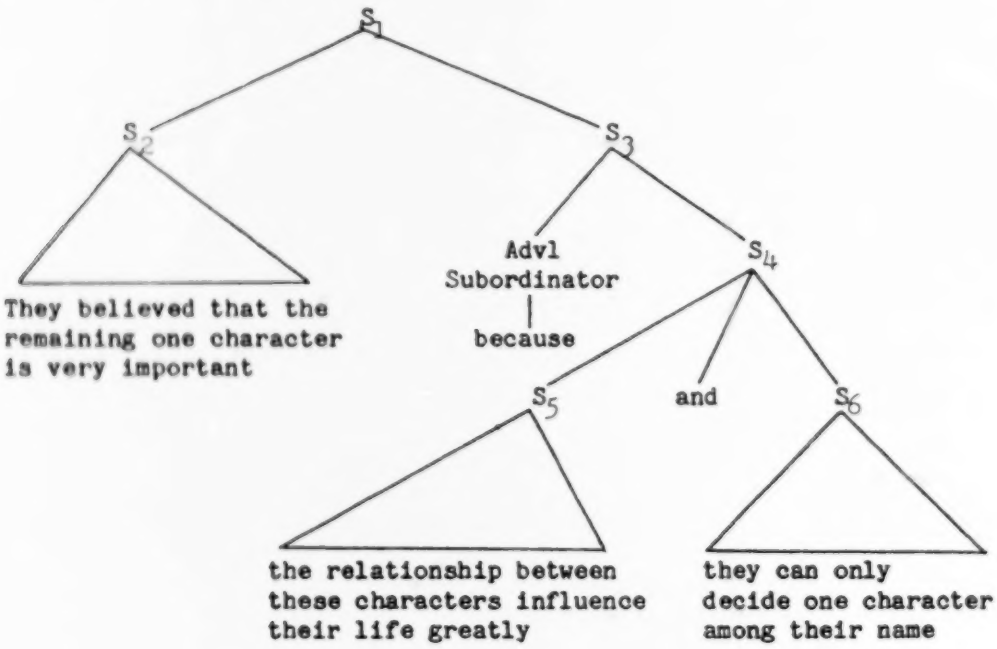
(2 T-units/ 1 sentence)

(13) Subordinate clause coordination



(1 T-unit / 1 sentence)

(14) They believed that the remaining one character is very important because the relationship between these characters influence their life greatly and they can decide only one character among their name. (L1 Korean, TOEFL 557)



Comparison of learners

Certainly a question which comes to mind is how a sentence-based analysis would account for the cases of excessive coordination (i.e., run-on sentences) which Hunt sought to avoid counting as legitimate when he developed the T-unit. We propose to use a coordination index. A coordination index gauges

the degree to which a learner achieves syntactic complexity through coordination (as opposed to subordination or embedding).

To calculate the coordination index, the number of clauses and sentences are tabulated. Because each sentence has at least one independent clause (or attempted independent clause in the case of a fragment), the number of sentences is subtracted from the total number of clauses to give the number of clauses which are connected in some way (by either coordination or subordination). Next, the number of independent-clause coordinations is counted. The number of independent-clause coordinations present in the language sample divided by the number of combined clauses yields the coordination index, which is in decimal form. Multiplied by 100 the coordination index gives the percentage of syntactic complexity achieved by independent-clause coordination. A sample calculation is given below:

(15) Calculating the Coordination Index:

Clauses	Sentences	Combinations = Clauses - Sentences	Coordinations	Coordination Index
47	22	47 - 22 = 25	6	6/25 = .24 x 100=24%

Our coordination index differs from Hunt's "main clause coordination index" in that the main clause coordination index measures the number of T-units per sentence (Hunt 1970, p. 189). Hunt's main clause coordination index gives a ratio of the total number of T-units to the total number of sentences. The coordination index proposed here takes into account only multiclausal sentences and reflects the frequency with which coordination is used by a learner relative to the total number of combinations produced in the language sample.

The coordination index quantitatively distinguishes between learners who employ coordination and those who employ subordination to attain syntactic complexity. Thus, if beginning learners (both L1 and L2) favor coordination over subordination (Monroe 1975), then the coordination index should be higher

for beginners than advanced learners. In this way, a sentence analysis with the coordination index should facilitate comparison between learners.

We analyzed 56 compositions written by English second language learners at seven levels of proficiency. The learners in levels 1-6 wrote on two similiar topics, "How have families changed in your country in the last thirty years?" and "How do you expect your life to be different from your grandparents'?" Data for the seventh level is taken from our forthcoming study of advanced learners (Bardovi-Harlig and Bofman, in press). Learners in the earlier study wrote on a variety of topics. The coordination index shows that the number of combinations of clauses resulting from coordination decreases with the proficiency of the learner. The T-unit analysis does not measure this type of development, but a sentence-based analysis with the coordination index does.

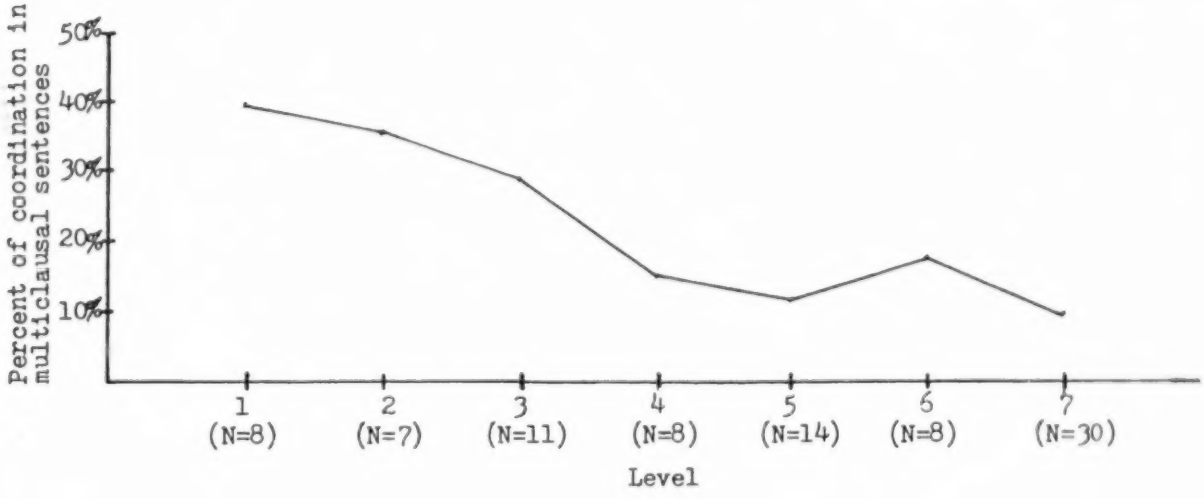


Figure 1. Coordination by Level (N=86)

Figure 1 shows a gradual decrease in the percentage of clauses which are combined by coordination. Learners who are less proficient rely on coordination to combine sentences to a greater extent than more advanced learners do.

Figure 2 compares the development of syntactic complexity across seven levels of proficiency in both T-units and sentences. In accordance with the widespread use of the term complex sentence by authors on a range of linguistic topics to mean a multi-clausal sentence exhibiting subordination (Quirk, Greenbaum, Leech, & Svartvik, 1985 on linguistic description; Clark & Clark, 1977, Foss & Hakes, 1978 and Rochester, 1973 on psycholinguistics; Feigenbaum, 1985 on pedagogical grammar; Pica, Young, Doughty, 1987 on input), we calculated syntactic complexity as the number of clauses per T-unit (Hunt's subordinate clause index (Hunt, 1965, 1970)). Both T-unit and sentence analyses show an increase in the number of clauses per syntactic unit as the learners' proficiency increases.

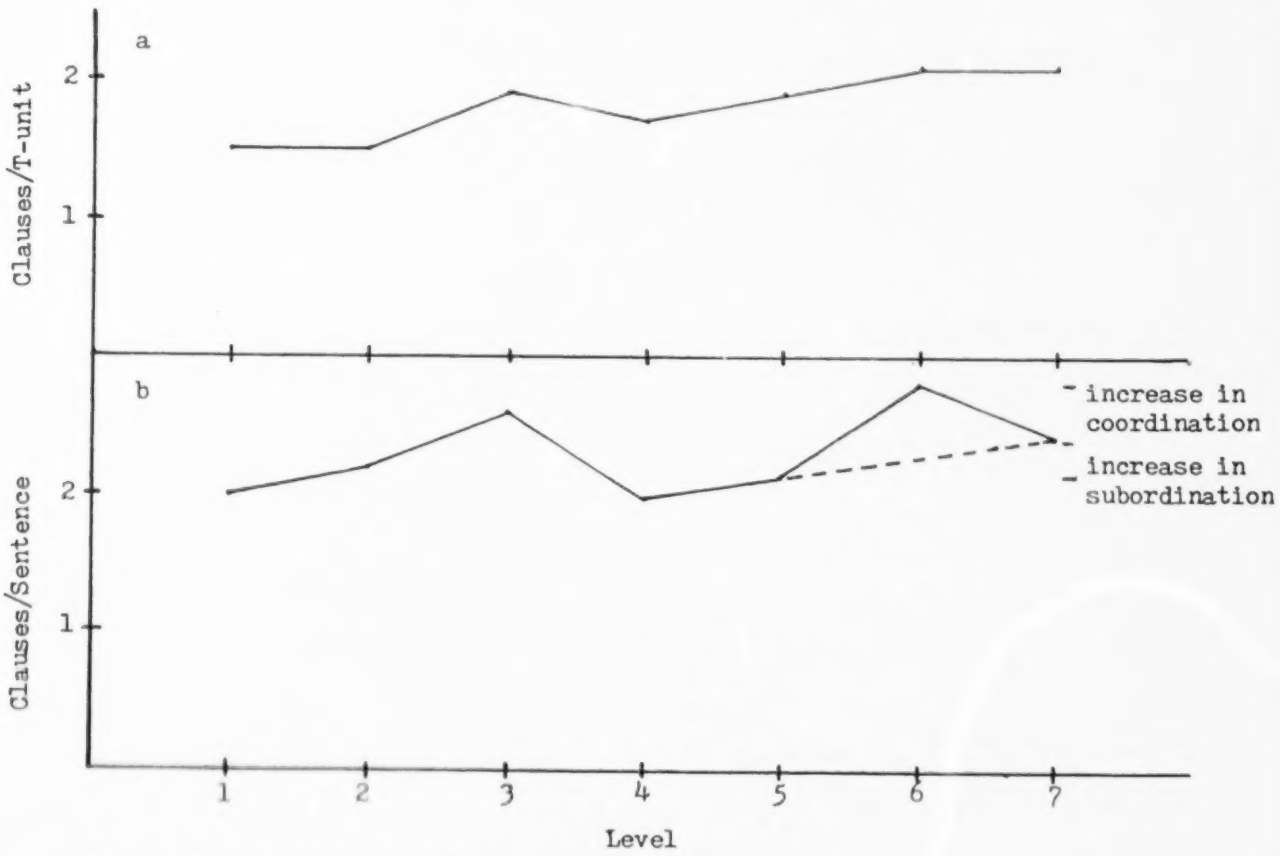


Figure 2. Complexity by Level (N=86)

At level 3 there seems to be a jump in subordination. Comparing Figure 2a with 2b, we see that 2b, the sentence analysis, also indicates an increase in coordination at level 6. The same increase is also pictured in Figure 1. At this time we don't know why level 6 shows this increase. This may be due to sample size. In contrast, the T-unit analysis, pictured in Figure 2a, gives no information on the development of coordination.

We should note that the coordination index as we have calculated it reflects only the number of instances of main clause coordination such as that diagrammed in example (12), but not coordination of embedded clauses such as that diagrammed in (14). We may want to suggest further modifications of the index as we consider more data.

Researcher-teacher communication

The use of the sentence, a learner-centered unit, encourages researcher-teacher dialogue in that both groups have access to the data in the same form. A unit which can be used by both researchers and teachers is not simply a luxury, but an essential ingredient in maintaining the researcher-teacher dialogue so important in our field. The unit should be useful and meaningful to both researchers and classroom teachers in second and foreign language acquisition and teaching in order to facilitate communication between the two groups. Stating research findings in sentence units uses terms most easily convertible to classroom use. It is precisely the sentence that is judged as a unit of development for evaluation, placement, and instruction by teachers in the classroom.

Conclusion

In conclusion, a sentence analysis has several advantages over a T-unit analysis for the description of syntactic complexity. A T-unit analysis completely eliminates full-clause coordination from any quantitative description of syntactic development, yet coordination is an important

recursive device in natural language. A sentence analysis allows knowledge of coordination to be represented in quantitative descriptions thereby more accurately reflecting learner knowledge. The more accurate description of developing syntactic knowledge facilitates comparison of learners at different levels of acquisition. In particular, the coordination index allows comparsion of the shift from the use of coordination by beginning learners to the use of embedding by more advanced learners. Finally, the use of the sentence, a learner-centered unit, encourages researcher-teacher dialogue in that both groups have access to the data in the same form. We believe that sentence-based analyses, whether they measure average length of sentences, average length of error-free sentences, or clauses per sentence, will prove fruitful in the analysis of written second language discourse.

Footnotes

*This paper was presented at the 22nd TESOL Convention, Chicago, Illinois, March 1988. We would like to thank Beverly Hartford and Don Siegel for comments on an earlier version of the paper. We would also like to thank Cheryl Engber for her assistance in coding.

[1]See Gaies (1980) for discussion of additional problems and limitations of the T-unit analysis.

[2]We follow Celce-Murcia and Larsen-Freeman's 1983 analysis of the adverbial subordinator.

References

Bardovi-Harlig, K. & Bofman T. (in press). Attainment of syntactic and morphological accuracy by advanced language learners. SSLA, 11 (1).

Beebe, L. (1983). Risk-taking and the language learner. In H. W. Seliger & M. H. Long (Eds.), Classroom oriented research (pp.39-65). Rowley, MA: Newbury House.

Celce-Murcia, M. & Larsen-Freeman, D. (1983). The grammar book. Rowley, MA: Newbury House.

Clark, H. H. & Clark, E. V. (1977). Psychology and language. New York: Harcourt Brace Jovanovich, Inc.

Feigenbaum, I. (1985). The grammar handbook. New York: Oxford University Press.

Foss, D. J. & Hakes, D. T. (1978). Psycholinguistics: An introduction to the psychology of language. Englewood Cliffs, NJ: Prentice-Hall.

Gaies, S. J. (1977). A comparison of the classroom language of ESL teachers and their speech among peers. Unpublished Ph.D. thesis, Indiana University.

Gaies, S. J. (1980). T-unit analysis in second language research: Applications, problems and limitations. TESOL Quarterly, 14, 53-60.

Hunt, K. W. (1965). Grammatical structures written at three grade levels. Research Report No. 3. Urbana, IL: National Council of Teachers of English.

Hunt, K. W. (1970). Recent measures in syntactic development. In M. Lester (Ed.), Readings in applied transformation grammar. New York: Holt, Rinehert and Winston.

Lakoff, R. (1971). If's, and's, and but's about conjunction. In C. J. Fillmore & D. T. Langendoen (Eds.), Studies in linguistic semantics (pp. 115-149). New York: Holt, Rinehart and Winston.

Larsen-Freeman, D. (1983). Assessing global second language proficiency. In H. W. Seliger & M. H. Long (Eds.), Classroom oriented research (pp.287-304). Rowley, MA: Newbury House.

Monroe, J. H. (1975). Measuring and enhancing syntactic fluency in French. The French review, 48, 1023-1031.

Pica, T., Young, R. & Doughty, C. (1987). The impact of interaction on comprehension. TESOL Quarterly, 21, 737-758.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). A comprehensive grammar of the English language. London: Longman.

Rochester, S. R. & Gill, J. (1973). Production of complex sentences in monologues and dialogues. Journal of verbal learning and verbal behavior, 12, 203-210.

Schmerling, S. (1976). Asymmetric conjunction and rules of conversation. In P. Cole & J. L. Morgan (Eds.), Speech acts (pp. 211-256). New York: Academic Press.